GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:58:25 ; Search time 16 Seconds
(without alignments)
623.987 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1410	32.6	820	6	US-10-467-657-4910	Sequence 4910, Ap
2	1395.5	32.2	804	7	US-11-070-080-18	Sequence 18, Appl
3	156	3.6	824	6	US-10-957-569-31	Sequence 31, Appl
4	156	3.6	824	7	US-11-097-589-29	Sequence 29, Appl
5	153.5	3.5	817	6	US-10-793-626-50	Sequence 50, Appl
6	153.5	3.5	817	6	US-10-793-626-1528	Sequence 1528, Ap
7	150.5	3.5	611	6	US-10-454-437-54	Sequence 54, Appl
8	150.5	3.5	697	6	US-10-485-517-202	Sequence 202, App
9	150.5	3.5	925	6	US-10-454-437-50	Sequence 50, Appl
10	144.5	3.3	716	6	US-10-131-826A-96	Sequence 96, Appl
11	143	3.3	220	7	US-11-040-595-4	Sequence 4, Appli
12	141.5	3.3	769	6	US-10-467-657-3280	Sequence 3280, Ap
13	141	3.3	655	6	US-10-467-657-3168	Sequence 3168, Ap
14	134.5	3.1	825	7	US-11-074-176-20	Sequence 20, Appl
15	134	3.1	389	6	US-10-821-234-1401	Sequence 1401, Ap
16	133.5	3.1	1786	7	US-11-196-400-3	Sequence 3, Appli
17	132	3.0	350	6	US-10-454-437-40	Sequence 40, Appl
18	131	3.0	709	7	US-11-074-176-158	Sequence 158, App
19	129.5	3.0	2665	7	US-11-124-368A-214	Sequence 214, App
20	129.5	3.0	2668	7	US-11-124-368A-215	Sequence 215, App
21	127.5	2.9	859	6	US-10-467-657-6084	Sequence 6084, Ap
22	126.5	2.9	488	6	US-10-485-517-307	Sequence 307, App
23	125.5	2.9	459	6	US-10-467-657-3092	Sequence 3092, Ap
24	124.5	2.9	440	6	US-10-821-234-1330	Sequence 1330, Ap
25	124	2.9	1960	7	US-11-069-834-48	Sequence 48, Appl
26	124	2.9	1960	7	US-11-069-834-50	Sequence 50, Appl
27	123.5	2.9	426	6	US-10-454-437-70	Sequence 70, Appl
28	123.5	2.9	426	6	US-10-454-437-72	Sequence 72, Appl
29	122	2.8	420	7	US-11-074-176-48	Sequence 48, Appl
30	120	2.8	406	6	US-10-821-234-1349	Sequence 1349, Ap
31	120	2.8	406	6	US-10-878-556A-192	Sequence 192, App
32	119	2.7	980	7	US-11-064-246-10	Sequence 10, Appl
33	119	2.7	980	7	US-11-169-041-141	Sequence 141, App
34	119	2.7	1652	6	US-10-995-561-663	Sequence 663, App
35	119	2.7	1938	6	US-10-995-561-661	Sequence 661, App
36	119	2.7	1938	6	US-10-995-561-662	Sequence 662, App
37	119	2.7	1954	6	US-10-995-561-660	Sequence 660, App
38	119	2.7	1972	6	US-10-995-561-664	Sequence 664, App
39	119	2.7	1972	6	US-10-995-561-666	Sequence 666, App
40	118	2.7	1268	6	US-10-995-561-918	Sequence 918, App
41	118	2.7	1268	6	US-10-995-561-919	Sequence 919, App
42	118	2.7	1268	6	US-10-995-561-920	Sequence 920, App
43	116	2.7	989	6	US-10-821-234-975	Sequence 975, App
44	115.5	2.7	2890	7	US-11-115-639-32	Sequence 32, Appl
45	115	2.7	1618	6	US-10-984-645-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-467-657-4910
; Sequence 4910, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4910
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4910

Query Match 32.6%; Score 1410; DB 6; Length 820;
Best Local Similarity 40.3%; Pred. No. 1.1e-94;
Matches 311; Conservative 153; Mismatches 256; Indels 52; Gaps 12;

```
Qy          65 DPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYPIAEV 124
               | | :      |:: || |  :||:    |     :|: || | ::: :       :: :
Db          60 DAAVEEPVATDLYQTGTVAQVLQVL--KLPDGTVKVLVEGLYRGRVLTIEDTGGLFVSHI 117

Qy         125 EQLDRLEEFPNTCKMREELGELSEQFYKYAVQLVEMLDMSVPA--VAKLRRLLDSLPREA 182
               |:   :   ||  :      |  || :||      |:  :||  :  :  : ::
Db         118 EAVVEEDTGGNT-DLEAVRRTLLAQFEQYA-----KLNKKIPAEIIGSINGIAEN---SR 168

Qy         183 LPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTR---KPKQDDD 239
               | | : : ::    :: |||:   : :|  :  :  |  ::: ::  :: |   | : :
Db         169 LTDTVAAHLQLKLAQRQQILEIPEIGKRMEFLLAKLESELDIMQAEKRIRGRVKRQMEKS 228

Qy         240 KRVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKK 299
               :|   :   :|  |   | :|||   | ::  ||  |: : | ::| : |: |:|:||
Db         229 QREYYLN--EQIKAIHKELGEEDE---NGELDALEAGIKKAGMTKEAEEKCLSELKKLKM 283

Qy         300 MPQSMPEYALTRNYLELMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAV 359
               ||    | : |||:: :: ||| | :    ||  | ::|| ||| :||:|:|:|||||
Db         284 MPPMSAESTVVRNYIDTLLGLPWKKKSRVSKDIAKAGLVLDADHYGLEKVKERILEYLAV 343
```

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:56:40 ; Search time 50 Seconds
(without alignments)
1408.794 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4331	100.0	852	2	US-09-741-150-2	Sequence 2, Appli
2	4331	100.0	852	2	US-10-160-187-2	Sequence 2, Appli
3	1869.5	43.2	884	2	US-09-741-150-4	Sequence 4, Appli
4	1869.5	43.2	884	2	US-10-160-187-4	Sequence 4, Appli
5	1745.5	40.3	810	2	US-09-902-540-11337	Sequence 11337, A
6	1407.5	32.5	820	2	US-09-252-991A-30204	Sequence 30204, A
7	1406.5	32.5	832	2	US-09-540-236-3056	Sequence 3056, Ap
8	1396.5	32.2	790	2	US-09-543-681A-6059	Sequence 6059, Ap
9	1390.5	32.1	822	2	US-09-328-352-5754	Sequence 5754, Ap
10	1378	31.8	845	2	US-09-489-039A-13590	Sequence 13590, A
11	1284	29.6	809	2	US-09-252-991A-29280	Sequence 29280, A
12	1251	28.9	937	1	US-08-253-155A-31	Sequence 31, Appl
13	1251	28.9	937	2	US-09-538-092-1092	Sequence 1092, Ap
14	1251	28.9	937	2	US-09-949-002-374	Sequence 374, App
15	1250	28.9	959	2	US-09-538-092-1091	Sequence 1091, Ap
16	1224	28.3	825	2	US-09-438-185A-29	Sequence 29, Appl
17	1123	25.9	838	2	US-09-248-796A-17648	Sequence 17648, A
18	976	22.5	783	2	US-09-248-796A-16773	Sequence 16773, A
19	886	20.5	440	2	US-09-198-452A-44	Sequence 44, Appl
20	407	9.4	562	2	US-09-949-002-426	Sequence 426, App
21	390	9.0	458	2	US-09-270-767-46541	Sequence 46541, A
22	353	8.2	157	2	US-09-861-451A-8	Sequence 8, Appli
23	336	7.8	424	2	US-09-198-452A-45	Sequence 45, Appl
24	327	7.6	191	2	US-09-270-767-62133	Sequence 62133, A
25	165	3.8	3433	2	US-09-091-501B-10	Sequence 10, Appl
26	165	3.8	3433	2	US-09-538-092-1136	Sequence 1136, Ap
27	162.5	3.8	900	2	US-09-248-796A-16848	Sequence 16848, A
28	161	3.7	761	2	US-09-328-352-7492	Sequence 7492, Ap
29	159	3.7	701	2	US-09-134-001C-3327	Sequence 3327, Ap
30	156.5	3.6	767	2	US-09-252-991A-31198	Sequence 31198, A
31	153.5	3.5	817	2	US-09-710-279-50	Sequence 50, Appl
32	153.5	3.5	817	2	US-09-710-279-1528	Sequence 1528, Ap
33	153.5	3.5	823	2	US-09-134-001C-4081	Sequence 4081, Ap
34	153.5	3.5	863	2	US-09-328-352-6730	Sequence 6730, Ap
35	152.5	3.5	726	2	US-09-252-991A-26767	Sequence 26767, A
36	152	3.5	422	2	US-09-134-001C-3034	Sequence 3034, Ap
37	151	3.5	335	2	US-09-134-000C-4044	Sequence 4044, Ap
38	150.5	3.5	611	2	US-09-602-777A-54	Sequence 54, Appl
39	150.5	3.5	697	2	US-08-816-177-2	Sequence 2, Appli
40	150.5	3.5	925	2	US-09-602-777A-50	Sequence 50, Appl
41	150	3.5	724	2	US-09-489-039A-12100	Sequence 12100, A
42	146.5	3.4	556	2	US-09-134-000C-3848	Sequence 3848, Ap
43	146	3.4	903	2	US-09-328-352-6479	Sequence 6479, Ap
44	145.5	3.4	689	2	US-09-134-000C-4499	Sequence 4499, Ap
45	145	3.3	546	2	US-09-489-039A-9627	Sequence 9627, Ap

ALIGNMENTS

RESULT 1
US-09-741-150-2
; Sequence 2, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Human
US-09-741-150-2

Query Match 100.0%; Score 4331; DB 2; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQLVRSRLLKGTSLQSTILGVI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQLVRSRLLKGTSLQSTILGVI 60

Qy         61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120

Qy        121 IAEVEQLDRLEEFPNTCKMREELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IAEVEQLDRLEEFPNTCKMREELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180

Qy        181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240

Qy        241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300

Qy        301 PQSMPEYALTRNYLELMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PQSMPEYALTRNYLELMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360

Qy        361 QLKNNLKGPILCFVGPPGVGKTSVGRSVAKTLGREFHRIALGGVCDQSDIRGHRRTYVGS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:48:40 ; Search time 259 Seconds
(without alignments)
2320.888 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4323	99.8	852	2	Q86WA8_HUMAN	Q86wa8 homo sapien
2	4319	99.7	852	2	Q8NCE9_HUMAN	Q8nce9 homo sapien
3	4292	99.1	852	2	Q5R6M5_PONPY	Q5r6m5 pongo pygma
4	4154	95.9	852	2	Q9DBN5_MOUSE	Q9dbn5 m mus muscu
5	3629	83.8	715	2	Q8N3B9_HUMAN	Q8n3b9 homo sapien
6	3376	77.9	840	2	Q5PQY6_BRARE	Q5pqy6 brachydanio
7	2972	68.6	581	2	Q96K43_HUMAN	Q96k43 homo sapien
8	2863	66.1	806	2	Q6TXI3_RAT	Q6txi3 rattus norv
9	2089	48.2	432	2	Q8BK80_MOUSE	Q8bk80 mus musculu
10	1890.5	43.7	874	2	Q4T681_TETNG	Q4t681 tetraodon n
11	1869.5	43.2	885	1	LONH1_MAIZE	P93647 zea mays (m
12	1861	43.0	884	2	Q94F60_9POAL	Q94f60 dichantheli
13	1850	42.7	886	2	Q6RS97_WHEAT	Q6rs97 triticum ae
14	1849	42.7	888	1	LONH1_ARATH	O64948 arabidopsis
15	1846.5	42.6	884	2	Q8GV57_ORYSA	Q8gv57 oryza sativ
16	1812.5	41.8	880	2	Q69SH2_ORYSA	Q69sh2 oryza sativ
17	1801	41.6	875	1	LONH1_SPIOL	O04979 spinacia ol
18	1784	41.2	843	2	Q4NPW6_9DELT	Q4npw6 anaeromyxob
19	1745.5	40.3	826	1	LON2_MYXXA	P36774 myxococcus
20	1596	36.9	819	2	Q747S2_GEOSL	Q747s2 geobacter s
21	1585.5	36.6	788	2	Q82V32_NITEU	Q82v32 nitrosomona
22	1573	36.3	775	2	Q5KWK1_GEOKA	Q5kwk1 geobacillus
23	1561	36.0	774	2	Q65GJ6_BACLD	Q65gj6 bacillus li
24	1559.5	36.0	776	2	Q817Q4_BACCR	Q817q4 bacillus ce
25	1546.5	35.7	796	2	Q89H21_BRAJA	Q89h21 bradyrhizob
26	1545.5	35.7	776	2	Q6HSS4_BACAN	Q6hss4 bacillus an
27	1545	35.7	557	2	Q8GT60_ARATH	Q8gt60 arabidopsis
28	1544.5	35.7	776	2	Q633X4_BACCZ	Q633x4 bacillus ce
29	1543.5	35.6	773	2	Q81LC1_BACAN	Q81lc1 bacillus an
30	1541.5	35.6	776	2	Q6HD56_BACHK	Q6hd56 bacillus th
31	1540	35.6	856	2	Q72CU2_DESVH	Q72cu2 desulfovibr
32	1535.5	35.5	773	2	Q4MS75_BACCE	Q4ms75 bacillus ce
33	1535.5	35.5	773	2	Q72ZV6_BACC1	Q72zv6 bacillus ce
34	1535	35.4	772	2	Q8CXB9_OCEIH	Q8cxb9 oceanobacil
35	1525.5	35.2	774	2	Q9K8F6_BACHD	Q9k8f6 bacillus ha
36	1525	35.2	783	2	Q7WMZ2_BORBR	Q7wmz2 bordetella
37	1525	35.2	783	2	Q7WBH0_BORPA	Q7wbh0 bordetella
38	1509	34.8	778	2	Q8RC23_THETN	Q8rc23 thermoanaer
39	1506	34.8	779	1	LON_BRECH	P36772 brevibacill
40	1498	34.6	775	2	Q5WEP1_BACSK	Q5wep1 bacillus cl
41	1497	34.6	774	1	LON1_BACSU	P37945 bacillus su
42	1493	34.5	779	2	Q84FG5_9BACL	Q84fg5 brevibacill
43	1482	34.2	768	2	Q74EN9_GEOSL	Q74en9 geobacter s
44	1477.5	34.1	932	2	Q4WVD9_ASPFU	Q4wvd9 aspergillus
45	1476.5	34.1	937	2	Q7SA85_NEUCR	Q7sa85 neurospora

ALIGNMENTS

RESULT 1
Q86WA8_HUMAN
ID Q86WA8_HUMAN PRELIMINARY; PRT; 852 AA.
AC Q86WA8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peroxisomal lon protease.
GN Name=LONP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA De Walque S., Van Veldhoven P.P.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ548761; CAD68987.1; -; mRNA.
DR EMBL; BC093912; AAH93912.1; -; mRNA.
DR EMBL; BC093910; AAH93910.1; -; mRNA.
DR HSSP; P08177; 1RR9.
DR MEROPS; S16.006; -.
DR Ensembl; ENSG00000102910; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:52:15 ; Search time 48 Seconds
(without alignments)
1707.848 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1869.5	43.2	885	1	T04321	endopeptidase La h
2	1801	41.6	875	2	T09142	endopeptidase La h
3	1745.5	40.3	827	1	A36895	endopeptidase La (
4	1525.5	35.2	774	2	B84031	ATP-dependent prot
5	1506	34.8	779	1	B42375	endopeptidase La (
6	1497	34.6	774	1	I40421	endopeptidase La (
7	1473.5	34.0	817	1	A49844	endopeptidase La (
8	1467	33.9	821	2	B75530	ATP-dependent prot
9	1453.5	33.6	813	2	G75331	ATP-dependent prot
10	1445.5	33.4	787	2	A72230	endopeptidase La (
11	1433	33.1	820	2	H81908	probable endopepti
12	1426	32.9	820	2	H81106	ATP-dependent prot
13	1406.5	32.5	798	2	G83420	Lon proteinase PA1
14	1403.5	32.4	786	2	G82141	ATP-dependent LA p
15	1400.5	32.3	799	2	C87492	ATP-dependent prot
16	1398	32.3	795	1	A70322	endopeptidase La (
17	1393	32.2	778	2	E97224	ATP-dependent Lon
18	1372.5	31.7	784	2	AE0558	Lon protease [impo
19	1372	31.7	823	2	AF3361	endopeptidase La (
20	1371.5	31.7	779	2	E97778	endopeptidase La (
21	1369.5	31.6	805	2	AH2731	ATP-dependent prot
22	1369.5	31.6	805	2	G97512	ATP-dependent prot
23	1368	31.6	848	2	C82712	ATP-dependent seri
24	1367.5	31.6	784	1	SUECLA	endopeptidase La (
25	1367	31.6	810	1	JC6045	endopeptidase La (
26	1365.5	31.5	784	1	A71704	endopeptidase La (
27	1364	31.5	803	1	A64070	endopeptidase La (
28	1362.5	31.5	784	2	E90690	endopeptidase La [
29	1358	31.4	799	2	A85541	hypothetical prote
30	1355	31.3	784	2	AC0383	endopeptidase La (
31	1351.5	31.2	786	2	A96956	ATP-dependent prot
32	1350	31.2	784	1	S47270	endopeptidase La (
33	1349	31.1	777	2	F84985	endopeptidase La (
34	1309.5	30.2	773	2	T27382	hypothetical prote
35	1301.5	30.1	964	1	T04325	probable ATP-depen
36	1293.5	29.9	831	1	A71825	endopeptidase La (
37	1289.5	29.8	881	1	B71316	endopeptidase La (
38	1286	29.7	799	2	F83549	probable ATP-depen
39	1282.5	29.6	835	1	C64692	endopeptidase La (
40	1259.5	29.1	791	2	D81310	endopeptidase La (
41	1251	28.9	937	1	S42366	endopeptidase La h
42	1244	28.7	819	2	E81681	proteinase, Lon fa
43	1242	28.7	819	1	C71527	endopeptidase La (
44	1238	28.6	962	1	S57342	endopeptidase La h
45	1235.5	28.5	813	1	D70176	endopeptidase La (

ALIGNMENTS

RESULT 1
T04321
endopeptidase La homolog (EC 3.4.21.-) LON1 precursor, mitochondrial - maize
N;Alternate names: ATP-dependent proteinase LON1; ATP-dependent serine proteinase La ho
N;Contains: adenosinetriphosphatase (EC 3.6.1.3)
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04321
R;Barakat, S.; Pearce, D.A.; Sherman, F.; Rapp, W.D.
Plant Mol. Biol. 37, 141-154, 1998
A;Title: Maize contains a Lon protease gene that can partially complement a yeast pim1-
A;Reference number: Z15282; MUID:98281582; PMID:9620272
A;Accession: T04321
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-885 <BAR>
A;Cross-references: UNIPROT:P93647; UNIPARC:UPI0000043401; EMBL:U85494; NID:g1816585; P
C;Genetics:
A;Gene: LON2
C;Function:
A;Description: serine proteinase
C;Superfamily: ATP-dependent Lon protease
C;Keywords: ATP; DNA binding; hydrolase; mitochondrial matrix; mitochondrion; molecular
F;409-416/Region: nucleotide-binding motif A (P-loop)
F;472-477/Region: nucleotide-binding motif B
F;781/Active site: Ser #status predicted

Query Match 43.2%; Score 1869.5; DB 1; Length 885;
Best Local Similarity 44.7%; Pred. No. 3.1e-98;
Matches 401; Conservative 161; Mismatches 269; Indels 67; Gaps 15;

```
Qy          3 SVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQLVRSRLLKGTSLQSTILGVIP- 61
              | ||:::|||| :|     |||||: :|    :  :::||   | :    :  ::||:|
Db          2 SDSPVELPSRLAVLPFRNKVLLPGAIVRIRCTNPSSVKLVEQELWQ--KEEKGLIGVLPV 59

Qy         62 ----------NTPDPASDA---------------QDL----PPL--HRIGTAALAVQV 88
                        :|   ||:                ||      |:  |  | || |:
Db         60 RDSEATAVGSLLSPGVGSDSGEGGSKVGGSAVESSKQDTKNGKEPIHWHSKGVAARALH- 118

Qy         89 VGSNWPKPH----YTLLITGLCRFQIVQVLKEKPYPIAEVEQLDRLEEFPNTCKMREELG 144
              :     ||     | ::: ||||| : ::    || :| | :||  :      :  :|
Db        119 LSRGVEKPSGRVTYIVVLEGLCRFSVQELSARGPYHVARVSRLDMTKTELEQAEQDPDLI 178

Qy        145 ELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPREALPDILTSIIRTSNKEKLQILDA 204
               || ||   |::|: :|:     | : : |||::|   | ||  :    | :|:| :||:
Db        179 ALSRQFKATAMELISVLEQKQKTVGRTKVLLDTVPVYRLADIFVASFEISFEEQLSMLDS 238

Qy        205 VSLEERFKMTIPLLVRQIEGL----KLLQKTRKPKQDDDKRVIAIRPIRRITHISGTLED 260
              | |: |      |: | :: :    |: ||          |  :  : :| |    |
Db        239 VHLKVRLSKATELVDRHLQSILVAEKITQKVEGQLSKSQKEFLLRQQMRAIKEELG---- 294

Qy        261 EDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKMPQSMPEYALTRNYLELMVEL 320
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:48:20 ; Search time 200 Seconds
(without alignments)
1871.754 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4331	100.0	852	5	ABG32363	Abg32363 Human ATP
2	4331	100.0	852	7	ADM56266	Adm56266 Human ATP
3	4331	100.0	852	9	ADV85873	Adv85873 Human pro
4	4323	99.8	852	4	AAB47562	Aab47562 Protease
5	4323	99.8	852	7	ADC10040	Adc10040 Human NOV
6	4319	99.7	852	4	AAM93785	Aam93785 Human pol
7	4319	99.7	852	8	ADL31771	Adl31771 Human pro
8	4316	99.7	852	5	ABP69443	Abp69443 Human pol
9	3520	81.3	693	7	ADJ70423	Adj70423 Human hea
10	2972	68.6	581	4	AAB95332	Aab95332 Human pro
11	2972	68.6	581	4	AAB99179	Aab99179 Human ATP
12	2957	68.3	581	8	ADR08963	Adr08963 Human pro
13	2921	67.4	571	4	AAB94288	Aab94288 Human pro
14	2201	50.8	432	4	AAB97167	Aab97167 Human ATP
15	2201	50.8	432	4	AAE11085	Aae11085 Human ATP
16	2165	50.0	433	6	ABU11611	Abu11611 Human MDD
17	1983.5	45.8	423	4	AAU17375	Aau17375 Novel sig
18	1983.5	45.8	423	7	ADB94083	Adb94083 Human nov
19	1974	45.6	391	5	ABB89248	Abb89248 Human pol
20	1869.5	43.2	884	7	ADM56268	Adm56268 Human ATP
21	1869.5	43.2	884	9	ADV85875	Adv85875 Zea mays
22	1849	42.7	888	8	ADN74069	Adn74069 Thale cre
23	1840.5	42.5	887	8	ADT60761	Adt60761 Plant pol
24	1745.5	40.3	810	9	ABM92138	Abm92138 M. xanthu
25	1745.5	40.3	827	8	ADS30399	Ads30399 Bacterial
26	1585.5	36.6	770	8	ADS30692	Ads30692 Bacterial
27	1557.5	36.0	769	8	ADS27589	Ads27589 Bacterial
28	1546.5	35.7	794	6	ABU17842	Abu17842 Protein e
29	1540.5	35.6	767	8	ADN27141	Adn27141 Bacterial
30	1540	35.6	786	8	ADN26262	Adn26262 Bacterial
31	1525.5	35.2	774	8	ADS28353	Ads28353 Bacterial
32	1497	34.6	774	8	ADS44794	Ads44794 Bacterial
33	1470.5	34.0	817	8	ADS30184	Ads30184 Bacterial
34	1468.5	33.9	771	8	ADS21365	Ads21365 Bacterial
35	1467	33.9	821	8	ADS44986	Ads44986 Bacterial
36	1461	33.7	777	8	ADN26487	Adn26487 Bacterial
37	1453.5	33.6	813	8	ADN17992	Adn17992 Bacterial
38	1450	33.5	788	8	ADS22145	Ads22145 Bacterial
39	1446.5	33.4	773	6	ABU24441	Abu24441 Protein e
40	1445.5	33.4	787	8	ADN20397	Adn20397 Bacterial
41	1433	33.1	820	6	ABU37980	Abu37980 Protein e
42	1426	32.9	789	6	ABU25481	Abu25481 Protein e
43	1426	32.9	820	8	ADP08282	Adp08282 Neisseria
44	1415.5	32.7	804	6	ABU39533	Abu39533 Protein e
45	1414.5	32.7	816	9	AEB41304	Aeb41304 L. pneumo

ALIGNMENTS

RESULT 1
ABG32363
ID ABG32363 standard; protein; 852 AA.
XX
AC ABG32363;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human ATP-dependent protease.
XX
KW Human; enzyme; ATP-dependent protease; developmental disorder; cancer;
KW retinoblastoma; melanotic melanoma; endometrial adenocarcinoma;
KW ovarian adenocarcinoma; schizophrenia.
XX
OS Homo sapiens.
XX
PN US2002081704-A1.
XX
PD 27-JUN-2002.
XX
PF 21-DEC-2000; 2000US-00741150.
XX
PR 22-NOV-2000; 2000US-0252410P.
XX
PA (GUEG/) GUEGLER K.
PA (WEBS/) WEBSTER M.
PA (YANC/) YAN C.
PA (SHAO/) SHAO W.
PA (KETC/) KETCHUM K A.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Guegler K, Webster M, Yan C, Shao W, Ketchum KA, Di Francesco V;
PI Beasley EM;
XX
DR WPI; 2002-635469/68.
DR N-PSDB; ABK90887, ABK90888.
XX
PT New human ATP-dependent peptides and encoding nucleic acids, useful for
PT diagnosing, preventing and/or treating disorders like cancer, e.g.
PT retinoblastomas, melanomas, endometrial and ovarian adenocarcinomas, and
PT schizophrenia.
XX
PS Claim 1; Fig 2; 119pp; English.
XX
CC The invention relates to a new isolated ATP-dependent protease (I) and
CC the nucleic acid encoding it. The peptides are useful in identifying

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: February 4, 2006, 13:57:30 ; Search time 177 Seconds
(without alignments)
2011.246 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 4331
Sequence: 1 MSSVSPIQIPSRLPLLLTHE.....AAFDGGFTVKTRPGLLNSKL 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4331	100.0	852	3	US-09-741-150-2	Sequence 2, Appli
2	4331	100.0	852	5	US-10-612-012-2	Sequence 2, Appli
3	4323	99.8	852	5	US-10-239-032-4	Sequence 4, Appli
4	3520	81.3	693	4	US-10-408-765A-2229	Sequence 2229, Ap
5	2201	50.8	432	4	US-10-156-239-44	Sequence 44, Appl
6	1983.5	45.8	423	3	US-09-764-868-940	Sequence 940, App
7	1974	45.6	391	4	US-10-264-237-1624	Sequence 1624, Ap
8	1869.5	43.2	884	3	US-09-741-150-4	Sequence 4, Appli
9	1869.5	43.2	884	5	US-10-612-012-4	Sequence 4, Appli
10	1869.5	43.2	885	4	US-10-156-239-47	Sequence 47, Appl
11	1840.5	42.5	887	5	US-10-739-930-10838	Sequence 10838, A
12	1745.5	40.3	827	4	US-10-369-493-19432	Sequence 19432, A
13	1657.5	38.3	850	4	US-10-437-963-156016	Sequence 156016,
14	1585.5	36.6	770	4	US-10-369-493-19725	Sequence 19725, A
15	1557.5	36.0	769	4	US-10-369-493-16622	Sequence 16622, A
16	1546.5	35.7	794	4	US-10-282-122A-45766	Sequence 45766, A
17	1540.5	35.6	767	4	US-10-369-493-9794	Sequence 9794,* Ap
18	1540	35.6	786	4	US-10-369-493-8915	Sequence 8915, Ap
19	1525.5	35.2	774	4	US-10-369-493-17386	Sequence 17386, A
20	1522	35.1	624	4	US-10-767-701-44549	Sequence 44549, A
21	1497	34.6	774	4	US-10-369-493-23224	Sequence 23224, A
22	1470.5	34.0	817	4	US-10-369-493-19217	Sequence 19217, A
23	1468.5	33.9	771	4	US-10-369-493-10398	Sequence 10398, A
24	1467	33.9	821	4	US-10-369-493-23416	Sequence 23416, A
25	1461	33.7	777	4	US-10-369-493-9140	Sequence 9140, Ap
26	1453.5	33.6	813	4	US-10-369-493-645	Sequence 645, App
27	1450	33.5	788	4	US-10-369-493-11178	Sequence 11178, A
28	1446.5	33.4	773	4	US-10-282-122A-52365	Sequence 52365, A
29	1445.5	33.4	787	4	US-10-369-493-3050	Sequence 3050, Ap
30	1433	33.1	820	4	US-10-282-122A-65904	Sequence 65904, A
31	1426	32.9	789	4	US-10-282-122A-53405	Sequence 53405, A
32	1415.5	32.7	804	4	US-10-282-122A-67457	Sequence 67457, A
33	1410.5	32.6	810	4	US-10-369-493-17875	Sequence 17875, A
34	1410	32.6	820	4	US-10-282-122A-65061	Sequence 65061, A
35	1409.5	32.5	798	4	US-10-369-493-13739	Sequence 13739, A
36	1408.5	32.5	799	4	US-10-369-493-20654	Sequence 20654, A
37	1406.5	32.5	798	4	US-10-282-122A-66328	Sequence 66328, A
38	1404.5	32.4	813	4	US-10-282-122A-48923	Sequence 48923, A
39	1403.5	32.4	786	4	US-10-282-122A-77359	Sequence 77359, A
40	1401.5	32.4	820	4	US-10-282-122A-63037	Sequence 63037, A
41	1401	32.3	817	4	US-10-282-122A-50881	Sequence 50881, A
42	1400.5	32.3	790	4	US-10-369-493-15757	Sequence 15757, A
43	1400.5	32.3	799	4	US-10-369-493-16934	Sequence 16934, A
44	1400.5	32.3	844	4	US-10-369-493-15389	Sequence 15389, A
45	1399.5	32.3	798	4	US-10-282-122A-67928	Sequence 67928, A

ALIGNMENTS

RESULT 1
US-09-741-150-2
; Sequence 2, Application US/09741150
; Publication No. US20020081704A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Human
US-09-741-150-2

Query Match 100.0%; Score 4331; DB 3; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQLVRSRLLKGTSLQSTILGVI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQLVRSRLLKGTSLQSTILGVI 60

Qy         61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120

Qy        121 IAEVEQLDRLEEFPNTCKMREELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 IAEVEQLDRLEEFPNTCKMREELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180

Qy        181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240

Qy        241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300

Qy        301 PQSMPEYALTRNYLELMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PQSMPEYALTRNYLELMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360

Qy        361 QLKNNLKGPILCFVGPPGVGKTSVGRSVAKTLGREFHRIALGGVCDQSDIRGHRRTYVGS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```